

①⑨ **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

⑫ **Offenlegungsschrift**  
⑩ **DE 198 47 261 A 1**

⑤① Int. Cl. 7:  
**G 06 K 9/62**

②① Aktenzeichen: 198 47 261.7  
②② Anmeldetag: 5. 10. 1998  
④③ Offenlegungstag: 6. 4. 2000

⑦① Anmelder:  
DCS Dialog Communication Systems AG, 10709  
Berlin, DE

⑦④ Vertreter:  
BOEHMERT & BOEHMERT, 28209 Bremen

⑦② Erfinder:  
Frischholz, Robert, Dr., 91058 Erlangen, DE

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ Verfahren und System zur Personenerkennung mit modellbasierter Gesichtsfindung

⑤⑦ Die Erfindung betrifft ein Verfahren und ein System zur Gesichtsfindung in einem binarisierten Bild durch Vergleichen der Punktegruppe des binarisierten Bildes mit der Punktegruppe eines Gesichtsmodells, bei dem die Punktegruppen des binarisierten Bildes und des Gesichtsmodells anhand des Hausdorff-Abstandes zwischen den Punkten der Punktegruppen verglichen werden und ein Gesicht in dem binarisierten Bild erkannt wird, wenn ein aus dem Hausdorff-Abstand abgeleitetes Maß einen Grenzwert unterschreitet.

**DE 198 47 261 A 1**

**DE 198 47 261 A 1**

## Beschreibung

Die Erfindung betrifft ein Verfahren und ein System zur Personenerkennung mit modellbasierter Gesichtsfindung.

Es gibt Verfahren und Systeme zur biometrischen Personenerkennung, die sich auf eine Kombination aus Gesichtserkennung, Spracherkennung und Lippenbewegungserkennung stützen. Der kritische Teil dabei ist die Gesichtsfindung. Die Erfindung betrifft ein Verfahren und ein System zur Lokalisierung des Gesichts einer Person in beliebigen Kamerabildern innerhalb eines solchen Systems zur biometrischen Personenerkennung.

Das grundlegende Element bei der Gesichtserkennung ist die Gesichtsfindung, d. h. die exakte Lokalisierung des Bildausschnittes, der dem menschlichen Gesicht entspricht. Im Stand der Technik vereinfachen einige Lösungsvorschläge diese Problematik, indem sie einen einheitlichen Hintergrund fordern, vor dem dann das Gesicht gefunden werden kann. Auch durch Hinzunahme der Bewegungsbildinformationen kann ein Gesicht erkannt werden, wobei hier unterstellt wird, daß nur der Bereich des Bildes, der sich im Laufe einer Bildfolge ändert, das Gesicht enthält.

Solche einfachen Ansätze genügen den heutigen Anforderungen an die Gesichtsfindung und -erkennung nicht mehr. Die Personenerkennung wird heute beispielsweise durch PC-Desktopkameras und andere mobile Kameras unterstützt, so daß die Gesichtsfindung auch dann robust funktionieren muß, wenn ein beliebiger Hintergrund oder sogar ein bewegtes Hintergrundbild vorliegt.

Die erfolgreichsten der heute bekannten Ansätze auf diesem Gebiet benutzen neuronale Netze. Diese Netze werden mit einer großen Anzahl von Gesichtsbeispielen trainiert, wobei bei dem Training als Gegenklasse eine noch größere Anzahl von "Nicht-Gesichtsbildern" verwendet wird, damit das neuronale Netz schließlich zwischen Gesicht und Hintergrund unterscheiden kann. Ein Nachteil dieser Methode ist neben der langen Rechenzeit die starke Abhängigkeit von Varianzen, wie Skalierung, Rotation und Verzerrung.

Die Erfindung hat daher zur Aufgabe, ein neues System und Verfahren zur Gesichtsfindung anzugeben, das robust ist und eine Gesichtsfindung in Echtzeit zu leisten vermag.

Zur Lösung dieser Aufgabe schlägt die Erfindung ein Verfahren mit den Merkmalen von Anspruch 8 vor.

Das erfindungsgemäße Verfahren und System sind modellbasiert. Das heißt, es wird ein binäres Bild eines menschlichen Gesichtes, oder eine sogenannte "Strichzeichnung" in einem entsprechenden Gesamtbild gesucht. Aus einem Originalbild wird also zunächst ein binarisiertes Bild, z. B. mittels Kantenextraktion, erzeugt, und dieses binarisierte Bild wird mit einem binären Gesichtsmodell verglichen, um das Bild des menschlichen Gesichtes in dem entsprechend binarisierten Gesamtbild zu suchen und zu finden. Es wird also nicht, wie bei den bisherigen Ansätzen des Standes der Technik, der Intensitätswert der einzelnen Kamerapixel verglichen oder nach Pixelveränderungen gesucht, sondern die Erfindung stützt sich auf eine Modellstruktur nach Art einer Strichzeichnung, die z. B. in Form einer Bitabbildung vorliegen kann.

Das Modell des Gesichtes oder die "Strichzeichnung", wird mit dem binarisierten Bild mit Hilfe einer modifizierten Version des sog. Hausdorff-Abstandes in zahlreichen Skalierungs-, Rotations- und Verzerrungsvarianten des Bildes verglichen. Die Anwendung des Hausdorff-Abstandes auf die Gesichtserkennung wurde z. B. von B. Takacs und H. Wechsler in "Face Recognition Using Binary Image Metrics", 2nd International Conference on Automatic Face and

geschrieben. Auf diese Veröffentlichung und die dortige Erläuterung des Hausdorff-Abstandes wird ausdrücklich Bezug genommen.

Die genannte Veröffentlichung beschreibt die Anwendung des Hausdorff-Abstandes zum Zwecke der Gesichtserkennung. Ausdrücklich wird darauf hingewiesen, daß die Gesichtsfindung mit einem gänzlich anderen Verfahren erreicht wird. Die Verwendung des Hausdorff-Abstandes zur Gesichtsfindung wurde im Stand der Technik nicht in Erwägung gezogen, u. a. weil dieser Algorithmus sehr viel Rechenzeit in Anspruch nimmt.

Dabei sollte man sich die grundsätzlichen Unterschiede in der Problematik der Gesichtsfindung und der Gesichtserkennung ins Gedächtnis rufen: Wenn das Gesicht in einem, insbesondere bewegten Bild einmal gefunden ist, kann dieses Gesichtsbild zur Erkennung bzw. Identifikation mit nahezu beliebigen Methoden mit einer Sammlung von Gesichtern aus einer Datenbank verglichen werden. Ist das fragliche Gesicht in der Datenbank enthalten, sind die Trefferraten im allgemeinen üblicherweise sehr hoch, bei etwa 99%. Das Schwierige bei der Personenerkennung ist jedoch der vorausgeschaltete Schritt, in einem beliebigen Bild zunächst das Gesicht zu finden und perfekt "auszuschneiden". Was bei der wortsinngemäßen Gegenüberstellung Gesichtserkennung/Gesichtsfindung also als geringfügiger Unterschied erscheinen mag, ist im Ergebnis für die Gesichts- und somit Personenerkennung entscheidend für die Güte des Ergebnisses.

Der Bildvergleich mithilfe des Hausdorff-Abstandes basiert auf folgenden Grundlagen:

Aus dem binarisierten Bild und dem Gesichtsmodell werden zwei Punktgruppen gebildet,

$$A = \{a_1, \dots, a_m\}$$

und

$$B = \{b_1, \dots, b_n\}$$

der Hausdorff-Abstand ist dann definiert durch  $H(A, B) = \max(h(A, B), h(B, A))$ , wobei

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|.$$

Das erfindungsgemäße System und Verfahren sind unempfindlich gegenüber den häufigsten Störeinflüssen, wie eine Drehung, unterschiedliche Skalierung oder Verzerrung des Bildes, weil diese bei dem Vergleich der Punktgruppen leicht berücksichtigt werden können: Auch ist für den Einsatz des erfindungsgemäßen Verfahrens kein langer Einlernvorgang nötig, anders als bei den neuronalen Netzen. Im Gegensatz zu den Ansätzen mit neuronalen Netzen müssen auch keine "Nicht-Gesichtsbilder" vorgegeben, gelernt oder auf andere Weise berücksichtigt werden. Das System erkennt ein Gesicht, wie der Mensch, aufgrund der Eigenschaften des Gesichtes selbst, nicht aufgrund der Eigenschaften des Hintergrundes, die somit nicht berücksichtigt werden müssen. Durch eine geeignete Ausbildung des Modells, der "Strichzeichnung", können gesuchte Besonderheiten (z. B. Brillenträger) schnell in dem Modell berücksichtigt werden.

Das erfindungsgemäße System und Verfahren können in der Zukunft in biometrischen Identifikationssystemen zur automatischen biometrischen Zugangskontrolle verwendet werden, bei denen häufig die Gesichtserkennung, Spracherkennung, Lippenbewegungserkennung, Retinaabtastung, etc. kombiniert werden. Mit Hilfe der Erfindung ist es möglich, für die Gesichtserkennung den exakten Gesichtsaus-

schnitt zu finden, die exakte Augenposition für eine Retinaabtastrung zu lokalisieren, die exakte Mundposition zur Berechnung der Lippenbewegungen zu lokalisieren und dergleichen.

Das erfindungsgemäße System und Verfahren können jedoch noch sehr viel universeller eingesetzt werden, indem z. B. durch Vorgabe entsprechender binärer Modelle nur Menschen mit bestimmten Gesichtszügen erkannt werden, um nach solchen Merkmalen wie Mimik, Rasse oder Geschlecht zu unterscheiden. Das erfindungsgemäße Verfahren und System sind nicht einmal auf die Gesichtsfindung beschränkt, weil das Modell, nach dem gesucht wird, auch eine Hand oder einen anderen oder mehrere Teile des menschlichen Körpers oder einer Sache umfassen könnte.

In ihren bevorzugten Ausführungsformen sieht die Erfindung für die Gesichtsfindung die Anwendung eines modifizierten Hausdorff-Abstandes vor, um den Rechenaufwand zu minimieren und so ein Ergebnis innerhalb einer vertretbaren Rechenzeit zu erhalten. Mit der derzeitigen Ausführungsform des modifizierten Hausdorff-Abstandes zur Gesichtsfindung kann etwa ein Gesichtsbild pro Sekunde detektiert werden.

Zusätzlich sieht die Erfindung eine neuartige Vorauswahl des Bildes durch eine spezielle Ausnutzung der Voronoi-Oberfläche vor, die auch zur Beschleunigung des Verfahrens beiträgt.

Das erfindungsgemäße System und Verfahren können sowohl als Software als auch als Hardwaremodule realisiert werden, wobei die modifizierten Algorithmen entweder programmiert oder in einer eigenen Hardware realisiert und somit wenigstens in der Hardwarelösung echtzeitfähig sind.

Die Erfindung ist im folgenden mit weiteren Einzelheiten in bezug auf die Zeichnungen näher erläutert. In den Figuren zeigen:

Fig. 1 ein Originalbild, das mit einer digitalen Kamera aufgenommen wurde;

Fig. 2 eine binarisierte Version des Originalbildes der Fig. 1;

Fig. 3 ein binäres Gesichtsmodell in Form einer Strichzeichnung; und

Fig. 4 die mit Hilfe des Gesichtsmodells der Fig. 3 in dem binarisierten Bild der Fig. 2 gefundene Position des Gesichtsmodells.

Fig. 1 zeigt das Originalbild einer Person mit Gesicht, Schulterpartie und Hintergrund, das mittels Kantenextraktion in das in Fig. 2 dargestellte binarisierte Bild umgewandelt wird. Hierzu werden die Kanten am Hell-Dunkel-Übergang in dem Originalbild dazu verwendet, eine Art Strichzeichnung des binarisierten Bildes gemäß Fig. 2 herzustellen. Gesucht wird nach dem zweidimensionalen Gesichtsmodell der Fig. 3 mit Hilfe des Hausdorff-Abstandes unter den unten näher erörterten Bedingungen.

Der allgemeine Hausdorff-Abstand bietet ein Mittel zur Bestimmung der Ähnlichkeit einer Punktgruppe zu einer anderen durch Untersuchung des Anteils der Punkte in der einen Gruppe, die in der Nähe von Punkten in der anderen Gruppe, oder umgekehrt, liegen. Es gibt zwei Parameter, um zu entscheiden, ob die zwei Punktgruppen einander ähneln oder nicht: (i) der maximale Abstand, den die Punkte voneinander entfernt sein können und bei dem sie gleichwohl noch als nah beieinander liegend betrachtet werden, und (ii) welcher Anteil der Punkte in einer Gruppe maximal diesen Abstand von den Punkten in der anderen Gruppe entfernt ist.

Die Gesichtsfindung mit Hilfe des Hausdorff-Abstandsmaßes unterscheidet sich von anderen Techniken, wie der binären Korrelation, weil es keine Paarung der Punkte in den beiden Gruppen gibt, die verglichen werden. Eine Erläuterung der mathematischen Grundlagen des Hausdorff-

Abstandes findet man im Internet unter der Adresse <http://www.cs.cornell.edu/Vision/hausdorff/hausmaich.html>. In diesem Dokument sind die Grundlagen des Hausdorff-Abstandes erläutert, auf die Bezug genommen wird.

Die zweidimensionale Bildabbildung der Fig. 3 dient also als ein Gesichtsmodell, das in dem binarisierten Bild der Fig. 2 lokalisiert werden soll, wobei geeignete zweidimensionale Transformationen und Skalierungen vorgenommen werden können.

Fig. 4 zeigt die beste Übereinstimmung des Modells der Fig. 3 mit dem binarisierten Bild der Fig. 2 und somit die gefundene Position des Modells in dem binären Bild in dem Sinne, daß in Fig. 4 der größte Anteil der binarisierten Kantenpunkte der Fig. 2 in der Nähe der Bildpunkte der Fig. 3 liegen. Bei Verwendung des Hausdorff-Abstandes findet man auch dann eine Übereinstimmung zwischen dem Gesichtsmodell und dem binarisierten Bild, wenn die entsprechenden binären Punkte nicht exakt übereinanderliegen.

Zur Realisierung dieses modellbasierten Gesichtsfindungsverfahrens werden bei einer bevorzugten Ausführungsform der Erfindung folgende Modifikationen des Hausdorff-Abstandes vorgenommen, um eine Gesichtsfindung in Echtzeit zu erreichen:

Die Erfindung verwendet einen hierarchischen Ansatz, bei dem zunächst das binarisierte Bild stark verkleinert wird, in dem verkleinerten binarisierten Bild nach einem entsprechend kleinen Gesichtsmodell gesucht wird; und wenn eine bestimmte Region als wahrscheinlicher Ort des gesuchten Gesichtes erkannt wird, wird diese Region und ihre Umgebung vergrößert, um den Suchvorgang mit einem entsprechend größeren Modellgesicht fortzusetzen, usw.

Bei dieser hierarchischen Suchweise werden unterschiedliche Modelle verwendet; d. h. bei dem stark verkleinerten binarisierten Bild wird z. B. ein Modell (Fig. 3) mit Schultern verwendet, um die Person zuverlässig zu erkennen. Je größer der Maßstab des binarisierten Bildes wird, desto feiner wird auch die Auflösung, und desto detaillierter wird das Modell für die Gesichtsfindung, so daß schließlich ein Modell verwendet werden kann, das z. B. nur noch Augen, Nase und/oder Mund enthält.

Auch die Kantenextraktion zur Herstellung des binarisierten Bildes der Fig. 2 kann an die jeweiligen Hierarchiestufen, in denen unterschiedlich feine Auflösungen benötigt werden, angepaßt werden. Hierzu sieht die Erfindung die Verwendung eines adaptiven Sobel-Filters vor.

Innerhalb der verschiedenen Hierarchiestufen können jeweils geeignete Rotationen des Bildes und/oder des Modells vorgenommen werden.

Zusätzlich sieht die Erfindung vorzugsweise eine Vorfilterung des binarisierten Bildes (Erosion, Dilatation, etc.) vor.

Ein weiteres wichtiges Merkmal der Erfindung ist die Modifikation des Hausdorff-Abstandes. Bei einer besonders bevorzugten Ausführungsform der Erfindung ist vorgesehen, mit einem modifizierten Hausdorff-Abstand zu arbeiten, bei dem nicht nur der mittlere Abstand aller minimalen Abstände zwischen Modell und Bild als Abstandsmaß verwendet wird, sondern bei dem der Mittelwert der ersten  $x\%$  ( $0 < x < 100$ ) aller minimalen Abstände als Grundlage der Berechnung des Hausdorff-Maßes dient, damit größere Abweichungen ("Ausreißer") nicht berücksichtigt werden und das Ergebnis verfälschen.

#### Patentansprüche

1. Verfahren zur Gesichtsfindung in einem binarisierten Bild durch Vergleichen der Punktgruppe des binarisierten Bildes mit der Punktgruppe eines Gesichts-

- modells, bei dem die Punktegruppen des binarisierten Bildes und des Gesichtsmodells anhand des Hausdorff-Abstandes zwischen den Punkten der Punktegruppen verglichen werden und ein Gesicht in dem binarisierten Bild erkannt wird, wenn ein aus dem Hausdorff-Abstand abgeleitetes Maß einen Grenzwert unterschreitet. 5
2. Verfahren nach Anspruch 1, bei dem das binarisierte Bild aus dem Originalbild mittels Kantenextraktion abgeleitet wird.
3. Verfahren nach Anspruch 1 oder 2, bei dem das binarisierte Bild zunächst in einem kleinen Maßstab mit einem entsprechend kleinen Gesichtsmodell verglichen wird, der Bereich des binarisierten Bildes, in dem ein Gesicht gefunden wurde, vergrößert und mit einem entsprechenden größeren Gesichtsmodell erneut verglichen wird und das Vergrößern und Vergleichen von binarisiertem Bildbereich und Gesichtsmodell ggf. wiederholt wird, bis das Gesicht in dem binarisierten Bild mit ausreichender Genauigkeit lokalisiert wurde. 10 15
4. Verfahren nach Anspruch 3, bei dem je nach Größe des binarisierten Bildes unterschiedliche Gesichtsmodelle mit unterschiedlicher Auflösung verwendet werden. 20
5. Verfahren nach Anspruch 3 oder 4, bei dem je nach Größe des binarisierten Bildes die Kantenextraktion für die Ableitung des binarisierten Bildes aus dem Originalbild mit unterschiedlicher Auflösung erfolgt. 25
6. Verfahren nach einem der Ansprüche 3 bis 5, bei dem je nach Größe des binarisierten Bildes unterschiedliche Rotationsstufen verwendet werden. 30
7. Verfahren nach einem der vorangehenden Ansprüche, bei dem das Hausdorff-Maß auf der Grundlage des Mittelwertes der kleinsten  $x\%$  aller minimalen Hausdorff-Abstände ermittelt wird, wobei  $0 < x < 100$ .
8. System zur Realisierung des Verfahrens nach einem der vorangehenden Ansprüche mit einer Rechenvorrichtung zur Berechnung des Hausdorff-Abstandes und des Hausdorff-Maßes auf der Grundlage der Punkte des binarisierten Bildes und des Gesichtsmodells. 35 40

---

Hierzu 1 Seite(n) Zeichnungen

---

45

50

55

60

65



Fig. 1

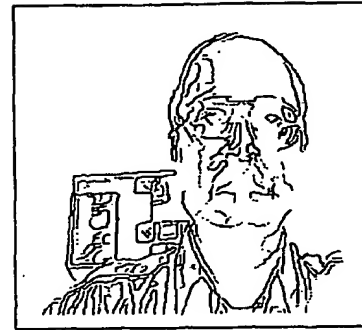


Fig. 2



Fig. 3

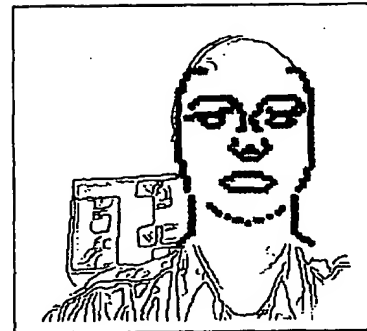


Fig. 4

- Leerseite -

*This Page Blank (uspto)*